Intro to Statistical Learning Theory
Exercise 2

1) Find the VC dimension of the following hypothesis spaces (prove your claim):

   a) Parity functions. $\mathcal{X} = \{0,1\}^n$, for any $S \subset [n]$ define $h_S(x) = \left(\sum_{i\in S} x_i\right) \mod 2$. $\mathcal{H} = \{h_S,\ S \subset [n]\}$.

   b) The set of rectangles in $\mathbb{R}^d$, i.e. $\mathcal{H} = \{h_{(c,b)}(x) = \mathbb{1}[\forall i, |x_i - c_i| \leq b_i],\ b, c \in \mathbb{R}^d\}$. We have seen in class $d = 2$.

   *c) The set of circles in $\mathbb{R}^2$, i.e. $\mathcal{H} = \{h_{(c,r)}(x) = \mathbb{1}[|x - c| \leq r],\ c \in \mathbb{R}^2,\ r > 0\}$

2) For $\mathcal{X} = \mathbb{R}$, define $\mathcal{H} = \{h_\theta(x) = \lceil \sin(\theta x) \rceil,\ \theta \in \mathbb{R}\}$ where we take $\lceil -1 \rceil = 0$. Prove that $VC(\mathcal{H}) = \infty$. Hint: prove and use the following lemma - if $x \in (0,1)$ has binary expansion $x = 0.x_1 x_2 ... x_m ....$ then for any natural number m, $\lceil \sin(2^m \pi x) \rceil = 1 - x_m$ provided that for some $k \geq m$ we have $x_k = 1$.

3) Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be binary hypothesis spaces over $\mathcal{X}$. define $d_i = VC(\mathcal{H})$, $d = \max(d_1, d_2)$ and assume $d \geq 3$.

   a) Prove that $VC(\mathcal{H}_1 \cup \mathcal{H}_2) \leq 2d + 1$

   b) Show and prove an upper bound on $VC(\mathcal{H}_1 \cdot \mathcal{H}_2)$, where $\mathcal{H}_1 \cdot \mathcal{H}_2$ is the class of all function of the form $h_1(x) \cdot h_2(x)$ when $h_i \in \mathcal{H}_i$.

   Hint: Use the bound we found on the growth function.

4) Structural Risk Minimization: If $\mathcal{H}$ has uniform convergence with complexity $\mathfrak{M}(\epsilon, \delta)$ we define the confidence $\epsilon(m, \delta) = \min_{\epsilon > 0}\{m > \mathfrak{M}(\epsilon, \delta)\}$, i.e. the best approximation error we can learn given m examples and probability δ. Prove the following theorem:

   Let p_n be a sequence of positive numbers such that $\sum_{n=1}^{\infty} p_n \leq 1$. Let $\mathcal{H} = \cup_{n=1}^{\infty} \mathcal{H}_n$ where $\mathcal{H}_n$ has uniform convergence with complexity $\mathfrak{M}_n(\epsilon, \delta)$ and confidence $\epsilon_n(m, \delta)$. For any distribution $\mathcal{D}$ we have with probability at least $1 - \delta$ over $S \sim \mathcal{D}^m$

$$\forall h \in \mathcal{H}, \quad L_{\mathcal{D}}(h) \leq L_S(h) + \min_{n: h \in \mathcal{H}} \epsilon_n(m, p_n \cdot \delta).$$

can you give a specific bound when $p_n = 2^{-n}$ and $VC(\mathcal{H}_n) = n$?